Radiometric Calibration Pipeline for the EO-1 Advanced Land Imager

Jenifer B. Evans and Herbert E. M. Viggh
MIT Lincoln Laboratory, Lexington, Massachusetts 02420

ABSTRACT

This paper covers the design and initial performance analysis of the radiometric calibration pipeline software for the EO-1 Advanced Land Imager (EO-1). The design and implementation of the software and of the radiometric calibration database are discussed. Radiometric calibration data were gathered in the laboratory during integration and test and were used for initializing the database. This database will be updated in the future with data collected on orbit. Initial performance results have been obtained after applying the radiometric calibration to imaging data collected during ALI integration and test and ground calibration. During the initial tests it was noted a few detectors leaked energy into other detectors, thereby requiring special calibration processing for the affected detectors. The handling of this anomaly is discussed, and the initial performance results of the calibration pipeline with the anomaly corrections will be shown.

Keywords: Radiometric, calibration, implementation, leaky pixel, performance assessment

1. INTRODUCTION

The Earth Observing-1 (EO-1) spacecraft mission is the first Earth observing mission under the National Aeronautics and Space Administration's (NASA) New Millenium Program. The goal of the mission is to validate new technologies that will contribute to the reduction in cost and increased capabilities for future land imaging missions. The Advanced Land Imager (ALI) is one of three instruments onboard the EO-1 spacecraft and was developed and built by MIT Lincoln Laboratory.

The ALI employs novel wide-angle optics and a highly integrated multispectral and panchromatic spectrometer. The focal plane of the ALI is partially populated with four sensor chip assemblies (SCA) and covers 3° by 1.625°. Operating in a pushbroom fashion at an orbit of 705 km, the ALI will provide one Landsat type panchromatic (PAN) and nine multispectral (MS) bands. These MS bands have been designed to mimic six Landsat MS bands with three additional bands covering 0.433-0.453, 0.845-0.890, and 1.20-1.30 μm. The ALI also contains wide-angle optics designed to provide a continuous 15° x 1.625° field of view for a fully populated focal plane with 30-meter resolution for the multispectral pixels and 10 meter resolution for the panchromatic pixels.

The ALI raw science data collected on orbit is downlinked and sent to NASA's Goddard Space Flight Center (GSFC) where it is converted to radiometrically calibrated data with engineering units of $Wcm^{-2}sr^{-1}$. This transform is performed by the radiometric calibration pipeline, designed and developed at MIT Lincoln Laboratory. This paper addresses the design and implementation of the pipeline, the initialization of the calibration database, initial performance assessment, and handling of leaky pixel data anomalies. Some aspects of this paper serve as a tutorial since there is not a large library of literature on the actual implementation of a calibration pipeline and database, and the relevant issues.

2. BACKGROUND

The ALI raw science data is downlinked, descrambled and arranged in imageable, band sequential order. This set of digital numbers, or raw sensor counts C_n, collected on orbit is referred to as the Level 0 science data. The radiometric calibration pipeline converts the Level 0 science data to the estimated in-band radiance L_λ with engineering units of $Wcm^{-2}sr^{-1}$. This radiometrically calibrated data is referred to as the Level 1R science data. The simple equation defining the transform is:

$$(1) \qquad \hat{L}_l = R_n(C_n - C_0)$$

where C_0 is the offset coefficient derived from the dark noise, and R_n is the response coefficient from the calibration database defining the relationship between the digital numbers and the estimated in-band radiance. The database of radiometric calibration response coefficients is derived from seven sources of radiometric calibration data. These seven sources are laboratory ground data, solar calibration data, lunar calibration data, deep space data, internal calibration lamps, dark data,

and flights over areas of known radiance. The main source for generating the coefficient database prior to launch is the laboratory ground calibration data.

During a typical on orbit data collection event (DCE), up to 30 seconds of science data are collected, followed by 8 seconds of internal calibration lamp data, and two seconds of dark data. These three sources of data form three datasets in a single Level 0 file. This file is the input to the radiometric calibration pipeline. The general software flow diagram is shown in Figure 1. The output of the calibration pipeline is a processing log and a single Level 1R file with three datasets corresponding to the radiometrically calibrated science data, the associated offset coefficients C_0, and the associated response coefficients R_n. The processing log contains warnings and general processing notes generated by the error checking functions or by the main calibration routine.

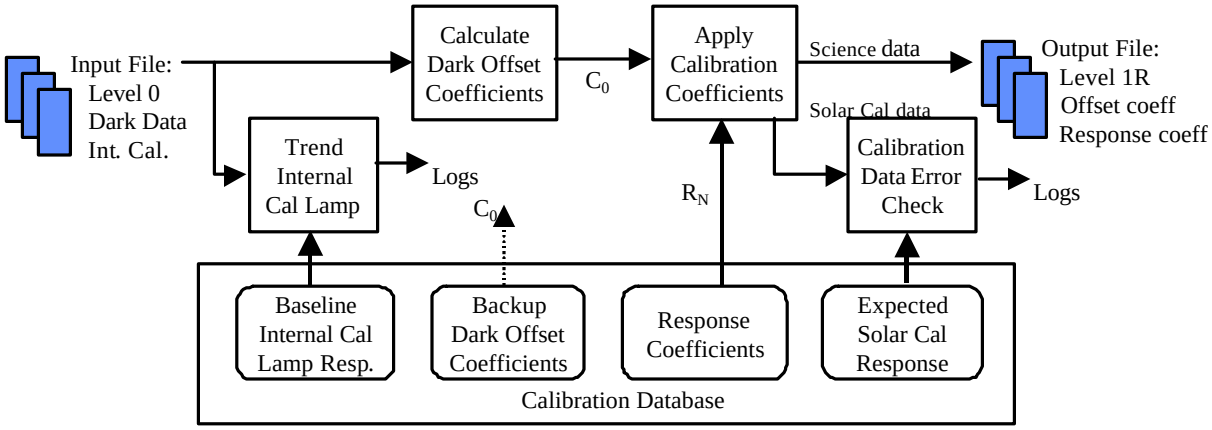


Figure 1. Calibration pipeline software flow diagram.

The calibration pipeline depends heavily on the calibration database, which contains four datasets. The most important dataset is the set of calibration response coefficients R_n, which are used regularly for the primary calibration. The details of this dataset will be examined in this paper. The other datasets within the calibration pipeline database are a default set of offset coefficients, a baseline internal calibration lamp dataset, and an expected solar response dataset. The offset coefficients, C_0, are generally computed for each DCE from the dark data included in the Level 0 file. However, if for some reason dark data is not available for a single DCE or all DCEs, e.g., the instrument cover no longer closes, a set of default offset coefficients exists in the database. The internal calibration lamp dataset that is collected along with each DCE is compared to the baseline internal calibration lamp dataset. Any significant change in response is noted in the calibration pipeline's processing log, and serves as a flag for detecting changes in the instrument. The expected solar calibration response dataset is used in a similar manner. The data collected during an on orbit solar calibration DCE will be radiometrically corrected using the calibration pipeline's offset and response coefficients. The expected radiometric solar response is compared to the measured response with any significant variation noted in the calibration pipeline's processing log.

3. PIPELINE DESIGN AND IMPLEMENTATION

3.1 Input/Output

The radiometric calibration pipeline is a stand alone, deliverable, software routine that receives Level 0 data as an input, performs the calibration, and outputs the Level 1R radiometrically calibrated data. Both the Level 0 data and the Level 1R data are available to various science teams performing data analysis in a distributed environment. Given the distributed work environment, the Hierarchical Data Format (HDF), created by the National Center for Supercomputing Applications (NCSA), was chosen for the file formats. HDF is self-describing, platform independent, and has supporting software for retrieval, storage, and visualization of the data, which makes the data easily readable by the distributed science teams.

In addition to containing the basic science data, the Level 0 and Level 1R data files also contain most of the ancillary data needed for the basic science data processing. This ancillary data is stored in HDF fields called attributes. Each single Level 0 or Level 1R data file contains multiple datasets. Each file has a set of file attributes pertaining to that datafile and all of its

datasets, e.g., file name, time of generation, and names of related files. Each dataset also has attributes with data parameters specific to that dataset, e.g., sensor used, dataset type, dimensions of dataset, integration time and frame rate. By using HDF and putting all of this information in a single, self-describing, platform independent file, any science team member is able to begin processing the data without further file format documentation. See Figure 2 for a depiction of a Level 0 file and content.

File Attributes		
Filename Data Product Level LZP Software version	Big or Little Endian Time of Generation Number Related Files	Related File 1 Related File 2 …
Dataset 1:	**Dataset 2:**	**Dataset 3:**
Science Data	Dark Data	Internal Lamp Data
Science Data Attrs	**Dark Data Attrs**	**Int. Lamp Data Attrs**
ALI Sensor Dataset Type Data Start Time Percent Missing Data …	ALI Sensor Dataset Type # along track pixels # bands …	ALI Sensor Dataset Type Integration Time Frame Rate …

Figure 2. Sample HDF Level 0 file contents and structure.

3.2 Calibration Pipeline Database

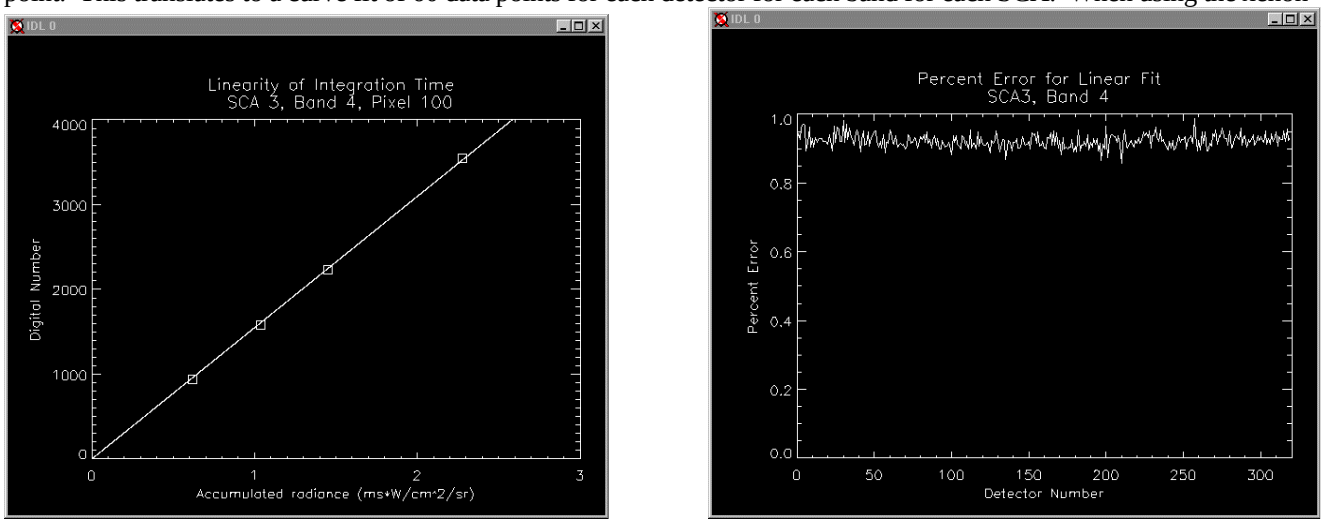
The radiometric calibration pipeline database is also an HDF file with multiple datasets. The primary dataset contains the response coefficients, with additional datasets for the default offset coefficients, the baseline internal calibration lamp response, and the expected solar calibration response. A typical database HDF file attribute is when the database was generated. Database dataset HDF attributes include the time when the baseline internal calibration lamp data was acquired, what files are the source of the default offset coefficients, and when the most recent on-orbit solar calibration data was acquired.

The main source of data for generating the calibration database prior to launch is the ground calibration data collected in the laboratory. For initializing the calibration response coefficient database, a total of 440 useable datasets were collected in the laboratory. The datasets include 96 xenon lamp datasets for calibrating the three shortest wavelength VNIR bands (MS 1', MS 1, and MS 2) and 344 halogen lamp datasets for the remaining VNIR (MS 3 MS 4, and MS 4') and SWIR (MS 5', MS 5, and MS 7.) Twelve different radiance levels were used, with the output of each level being monitored and recorded by a spectroradiometer. Three hundred and twelve of the datasets were recorded with the focal plane at the nominal operating temperature of 220 degree Kelvin, 64 datasets at 215 degree Kelvin, and 64 datasets at 225 degree Kelvin. The fourth variable was the integration time. The ALI integration times vary from 0 to 15, which correspond to 0.81 ms to 4.86 ms in steps of 0.27 ms. An integration time of 12 resulted in saturation for most bands at most radiance levels which completely exercised the full dynamic range of the focal plane.

A first order analysis of the laboratory ground data showed the system response to be relatively linear with respect to integration time. Figure 3a shows the best linear fit for a single detector with four different integration times and all other parameters (lamps, radiance levels, temperature) held constant. The percent error for a linear fit for this detector is 0.86%. Figure 3b shows the percent errors for all 320 detectors for a typical single band and single SCA. All errors are less than 1.0%.

After determining that the radiometric response is essentially linear with respect to integration time, it becomes feasible to use all the radiance levels and all the integration times, at a single temperature and lamp configuration to determine the best curve fit for each detector's calibration response coefficients. When using the halogen lamps at a focal plane temperature of

220 degree Kelvin, each detector has data for 5 integration times and 6 radiance levels with two sets of data collected at each point. This translates to a curve fit of 60 data points for each detector for each band for each SCA. When using the xenon



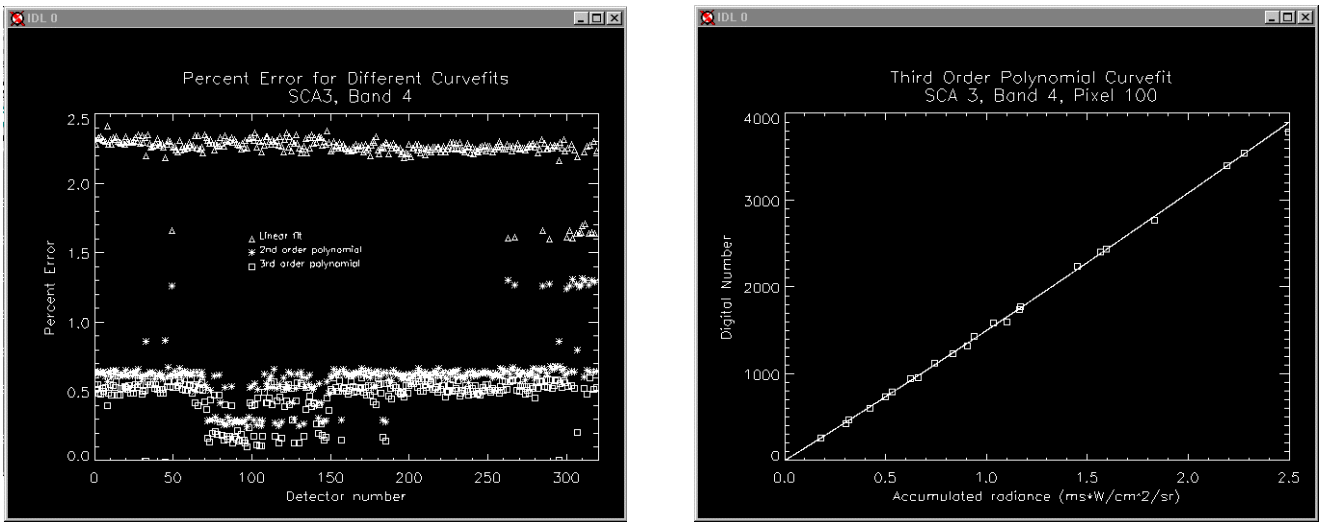
(a) (b)

Figure 3. Relationship between radiometric response and integration time. a.) Response is linear. b.) Errors less than 1% for linear fit.

lamps for the 3 shortest wavelengths, each detector has data for 2 integration times at 12 radiance levels with 2 sets collected at each point. This translates to a curve fit of 48 data points for each detector for these 3 bands for each SCA. In addition to fitting the curves to the measured data, all curves must also pass through (0,0). All of the data has been offset corrected with dark data recorded in conjunction with each dataset. If there is no flux, and offset correction has been performed, the digital number should be 0.

Ias a first cut, a linear fit was applied to the data. Average percent errors between the 60 measured data points and the curve fit were approximately 2 – 3% for each band. The specification for relative radiometric accuracy is 2% so a linear fit is not acceptable. A second order polynomial fit and a third order polynomial fit were also applied. Figure 4a shows the average percent errors for each of these fits for MS Band 4, on SCA 3. There is significant improvement over the linear fit, and slight improvement in general for the third order fit over the second order fit. Higher order polynomials did not show significant improvement until a sixth order fit. However, there is no obvious scientific reason to justify such a fit, and most likely the improvement is due to enough degrees of freedom in the curve to follow the noise and uncertainty in the measurements. For the majority of pixels, a second order or third order polynomial, whichever has the lowest percent error, is used for generating the calibration response coefficients. Figure 4b shows a third order polynomial curve for detector 100, MS Band 4, SCA 3 along with the measured data points used to derive the curve. Error bars are approximately the same height as the symbols shown in the plot.

The second- or third-order polynomial response curve is the data that needs to be represented in the calibration response coefficient database. In equations and diagrams, these database values are generally represented by R_n. If the radiometric response of all of the detectors had been strictly linear, this database could have been implemented as a collection of single response coefficients, R_n, one coefficient for each detector for each band for each SCA. But that wasn't the case. If all of the detectors showed a second order response relationship, this database could consist of two coefficients, R_{n0} and R_{n1}, for each detector for each band for each SCA. However as just shown, it was determined different order polynomials were appropriate for different detectors. Additionally, although not discussed in this paper, there are a few detectors requiring nonlinear, heuristic response functions. Given this situation, the primary response coefficient dataset was implemented as a lookup table with 4096 values for each detector for each band for each SCA. This is computationally faster and more flexible than traditional coefficient databases, at the cost of requiring more disk space and storage on the system running the calibration pipeline.



(a) (b)

Figure 4. Curvefit for calibration response coefficient. a.) Errors for linear, second-order, and third-order fits. b.) Typical third-order polynomial curvefit.

4. PERFORMANCE ASSESSMENT

The task of performance assessment can be broken down into two categories: performance assessment of the ALI system, including stability of the system and focal plane anomalies; and performance assessment of the radiometric calibration pipeline in terms of the absolute and relative accuracy of the radiometric calibration.

4.1 System level performance assessment

Performance assessment of the system includes monitoring the system for stability. The radiometric calibration pipeline includes a function for monitoring the internal calibration lamp data on a regular basis. The internal calibration lamp data was baselined prior to launch and will be trended over the course of the mission. The calibration pipeline will compare each new set of internal calibration data to the baseline. If there is significant variation, a warning will be included in the calibration pipeline processing logs. The importance of the internal calibration lamp data to the radiometric calibration pipeline is not for absolute radiometric calibration but for stability monitoring and overall performance assessment of the system.

Additionally the mean and standard deviation of the dark data, which is collected with every DCE, is continuously being calculated when determining the offset coefficients C_0. These numbers have been monitored throughout the I&T process, and will continue to be monitored on orbit to assure the system hasn't changed.

During ground performance assessment of the system as a whole, two focal plane anomalies were discovered after application of the radiometric calibration pipeline. Leaky pixels were found on SCA 3, Band 3, detector/pixel 225 and on SCA 4, Band 2, detector/pixel 190. The focal plane electronics of the ALI are such that all even numbered detectors on a single SCA share a gate, and all odd numbered detectors share a gate. So an odd numbered leaky pixel such as pixel 225 manifests itself by leaking energy into all other odd numbered detectors or pixels on that band on that SCA and artifically raising the counts, C_n, seen on those detectors. Figure 5 is from the integration and test (I&T) imaging configuration using the Air Force image test reticle. The effect of Pixel 225 being illuminated and leaking into the other odd-numbered pixels is visible to the eye.

Ideally, the leaky pixel could be completely characterized during ground calibration by masking the leaky pixel, performing the radiometric calibration of all other detectors when unaffected by the leaker, then unmasking the leaky pixel and repeating the process. The effect of the leaky pixel on each detector would be precisely known. Unfortunately, the I&T imaging

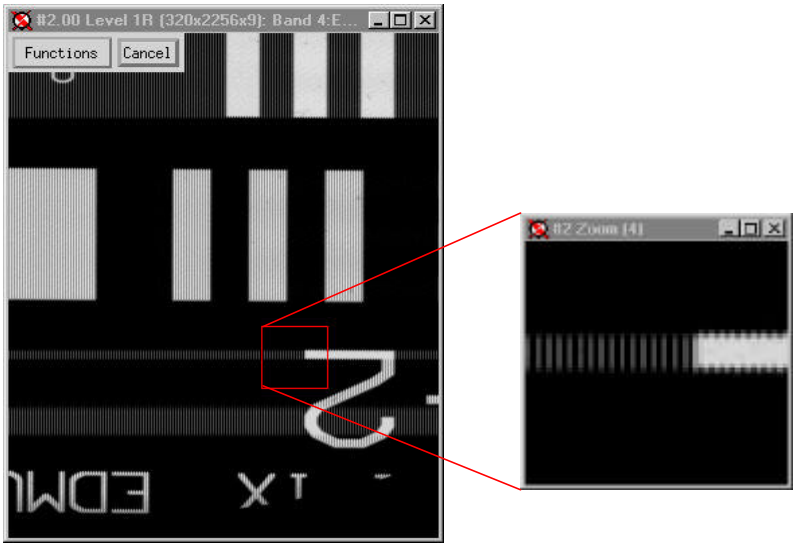


Figure 5. Effect of leaky pixel.

configuration which is necessary for masking of pixels and scanning of knife edges, does not support use of the halogen lamps and integrating sphere used in the radiometric calibration process which exercises the full dynamic range of the focal plane. The laboratory I&T imaging configuration does allow for limited dynamic range data collection. This data can be used to estimate the appropriate compensation for the leaky pixel and extrapolated to the higher radiances, and hopefully more finely adjusted after additional data is collected on orbit. The following discussion focuses on pixel 225 on SCA 3, Band 3, but also applies to pixel 190 on SCA 4, Band 2.

Given the limited available data, the first assumption made is that the leaky pixel, pixel 225, can be calibrated the same as all other normal pixels, thereby allowing us to determine L_l, the radiance on the leaker. The second assumption made is that the effect of the leaky pixel can be described by a linear equation. This equation describes the digital number or counts measured on pixel n, C_n, as a function of the radiance on pixel n, L_n, plus a percentage β_n of the radiance measured on the leaky pixel, L_l. The equation is written as:

$$C_n = a_n L_n + b_n L_l \qquad (2)$$

The notation defined in Table 1 is used in the following discussion. The first subscript to C denotes whether a pixel n is illuminated or dark, and the second subscript denotes whether the leaky pixel *l* is illuminated or dark. I is illuminated and Z is dark. Therefore:

Notation	Pixel n	Pixel l
$C_{I\,I}$	Illuminated	Illuminated
$C_{I\,Z}$	Illuminated	Dark
$C_{Z\,I}$	Dark	Illuminated
$C_{Z\,Z}$	Dark	Dark

Table 1. Notation for counts on detector n.

The other assumptions made are that L_n and L_l are either 0, or illuminated at a constant radiance of the same value. And since all processed data has been offset corrected, C_{ZZ} is considered to be 0. Given the above notations and assumptions, the equations are:

(3)

if

$$L_l = 0$$

then

$$C_n = \alpha_n L_n$$

$$C_{IZ} = \alpha_n L_n$$

$$C_{ZZ} = \alpha_n L_n = 0$$

Essentially, if the leaker is not illuminated, the counts on pixel n are purely a function of the radiance on pixel n and it's radiometric coefficient, α_n. If the leaker is iluminated, the following equations apply:

(4)

$$C_{II} = \alpha_n L_n + \beta_n L_l$$

$$C_{II} = C_{IZ} + \beta_n L_l$$

$$\beta_n = \frac{C_{II} - C_{IZ}}{L_l}$$

and

(5)

$$C_{ZI} = \alpha_n L_n + \beta_n L_l$$

$$C_{ZI} = C_{ZZ} + \beta_n L_l$$

$$\beta_n = \frac{C_{ZI} - C_{ZZ}}{L_l}$$

Essentially, the coefficient β_n describing the amount of energy leaking from pixel l to pixel n can be estimated by determining the difference in counts with and without the leaky pixel illuminated, divided by the radiance of the leaky pixel.

Data was collected with the I&T imaging configuration for the above 4 cases: pixel 225 dark, remaining pixels illuminated; pixel 225 illuminated, remaining pixels dark; all pixels illuminated; and all pixels dark. Given these four datasets, βn can be estimated with both equations (4) and (5). If the first assumption on linearity is true, the estimates of β_n should be the same. Figure 6 shows the estimates of β_n from the two equations. They are not the same, although they are within 10% of each other. The difference in the estimates of β_n indicates the amount of energy leaked from pixel 225 into the other pixels is not only a function of the radiance on pixel 225, but also a function of the radiance on the individual pixels. Initial analysis shows the higher the radiance on pixel n, the lesser the impact of leaky pixel 225 on the measured counts C_n.

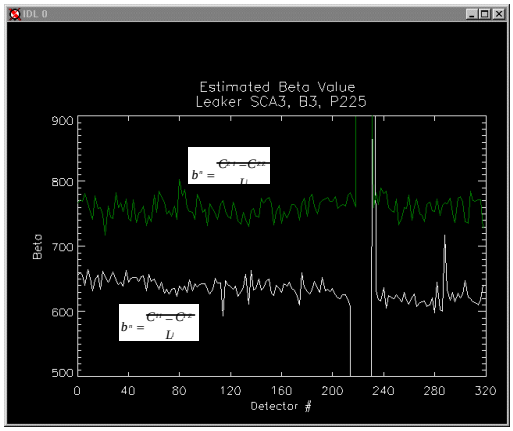
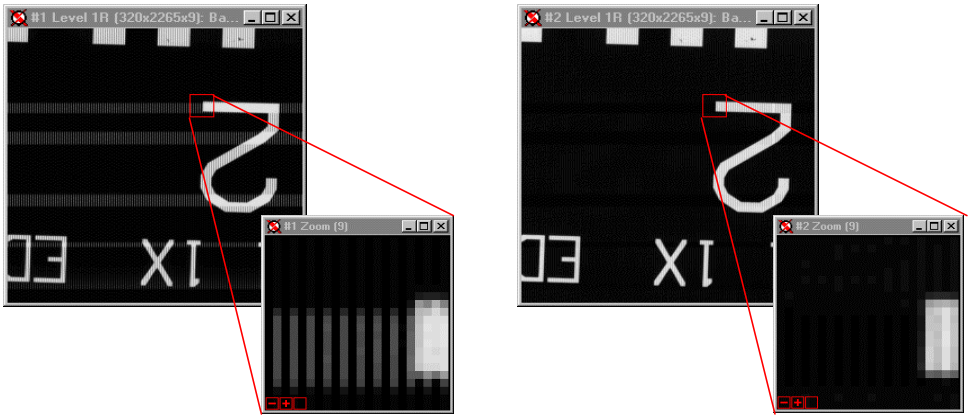


Figure 6. Two estimates for β_n using equations (4) and (5).

Without more laboratory I&T data, it is not possible at this time to more accurately determine the best correction for the leaky pixel. The first order linear approximation for βn can be used initially for the radiometric calibration, with attempts to more accurately determine the best correction after receiving more data collected on orbit. The effect of the current correction factor can be seen in Figure 7. To the naked eye, the correction is sufficient. Statistically, there is room for improvement.



(a) (b)

Figure 7. Affect of applying leaky pixel correction factor β_n. a.) Without correction. b.) With correction.

4.2 Calibration pipeline performance assessment

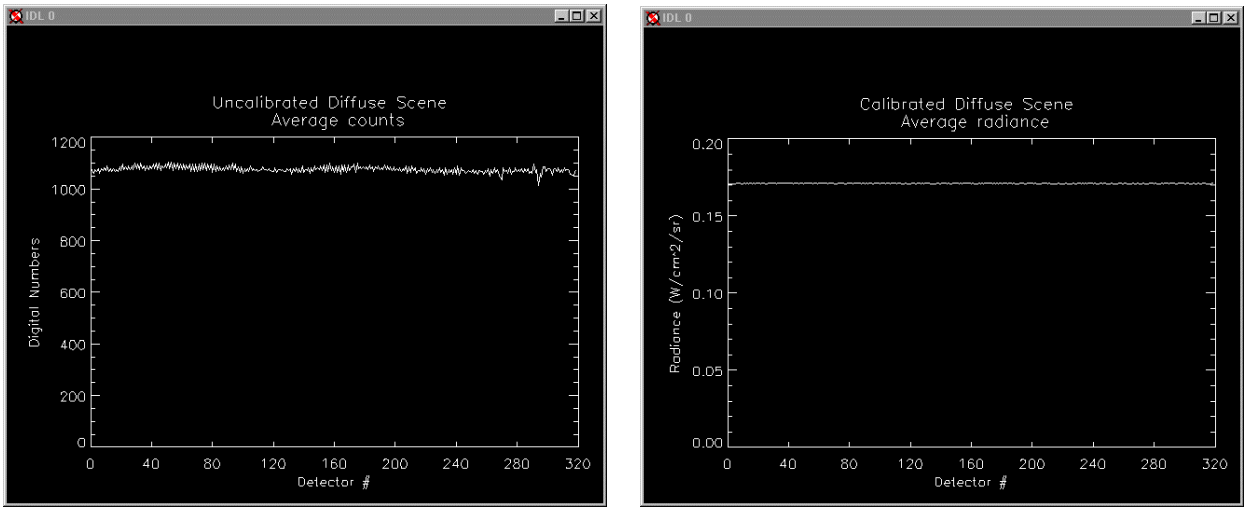
The majority of the performance assessment for the radiometric calibration pipeline will occur after launch. Seven sources of radiometric calibration data were described earlier. A number of these sources are data that will be acquired on orbit and will be used to assess the relative and absolute accuracy of the radiometric calibration. For example, lunar calibrations, deep-space calibrations, solar calibration, and flights over areas of known radiance all will result in data with a known or consistent radiometric response. If evaluation of this data reveals differences between the measured and the expected response, each of these sources of data will be used in conjunction with the ground data, internal calibration lamp data, and the dark data to determine the proper response. Possible results from further analysis include a determination of a need to modify the calibration coefficients, a determination the solar diffuser has begun to degrade, a determination the focal plane has changed, a determination the mirrors have become contaminated, or some other conclusion. Because of the complexity of the instrument, the decision and any resulting modifications are not automated.

The only pre-launch performance assessment of the calibration response coefficients is done by measuring the standard deviation across the focal plane of radiometrically calibrated data from a diffuse scene. However, since the only available uniform, diffuse data is the same data as used to generate the calibration response coefficient database, the consistent response seen is not surprising. Figure 8 shows the graphical output of this exercise in conjunction with the uncalibrated output. The software developed for this test will be used in the future on science data collected on orbit as part of the overall performance assessment of the radiometric calibration pipeline and the ALI.

5. SUMMARY.

In summary, the radiometric calibration pipeline developed for the ALI is operational and has been initialized with ground calibration data. The implementation utilizes HDF files for the input and output thereby making it easy to include ancillary data with the data files for use by a distributed team of scientists. At this point in time, the calibration response coefficients are represented by second- or third-order polynomials. After collecting more data on orbit, these response coefficients may change but will most likely remain polynomials of the same or similar degree.

The only significant focal plane anomalies detected during radiometric calibration of the I&T data are the two leaky pixels. It appears the leak is nonlinear, however, the existing I&T and ground calibration data is not sufficient to fully determine the best nonlinear model. Instead the leak has been represented with a linear model and initial indications are the linear representation may be sufficient. Once on orbit after additional data is collected, a more accurate model may be developed if necessary.



(a) (b)

Figure 8. (a) Uncalibrated output and (b) radiometrically calibrated output from a diffuse source.

ACKNOWLEDGMENTS

This work sponsored by NASA under Air Force Contract No. F19628-95-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the United States Air Force.